Identification of high-pressure phases III and IV in hydrogen: Simulating Raman spectra using molecular dynamics

Identification of high-pressure phases III and IV in hydrogen: Simulating Raman spectra using molecular dynamics

Ioan B. Magdău and Graeme J. Ackland
CSEC, SUPA, School of Physics and Astronomy, The University of Edinburgh, Edinburgh EH9 3JZ, United Kingdom
(Received 21 February 2013; published 23 May 2013)

We present a technique for extracting Raman intensities from *ab initio* molecular dynamics (MD) simulations at high temperature. The method is applied to the highly anharmonic case of dense hydrogen up to 500 K for pressures ranging from 180 to 300 GPa. On heating or pressurizing we find first-order phase transitions under the experimental conditions of the phase III-IV boundary. At even higher pressures, close to 350 GPa, we find a second phase transformation to the previously proposed *Cmca*-4. Our method enables, for the first time, a direct comparison of Raman vibrons between theory and experiment at finite temperature. This turns out to provide excellent discrimination between subtly different structures found in MD. We find candidate structures whose Raman spectra are in good agreement with experiment. The new phase obtained in high-temperature simulations adopts a dynamic, simple hexagonal structure with three layer types: freely rotating hydrogen molecules, static hexagonal trimers, and rotating hexagonal trimers. We show that previously calculated structures for phase IV are inconsistent with experiment, and their appearance in simulation is due to finite-size effects.

DOI: 10.1103/PhysRevB.87.174110 PACS number(s): 61.50.Ah, 61.66.Bi, 62.50.−p, 67.80.ff

I. INTRODUCTION

There have been some notable recent successes of using total energy calculations based on density functional theory (DFT) to calculate expected signals from candidate structures, for comparison with inconclusive experimental data. Agreement provides validation of the DFT structure, and this combined approach can yield more information, with higher reliability, than either technique alone.

Raman spectroscopy provides one such experimental probe, applicable in extreme conditions but providing insufficient data to determine crystal structure or identification of the vibrational mode.[1,2] Reliable calculation of Raman frequencies and intensities of mechanically stable structures can be obtained using density functional perturbation theory (DFPT)[3,4] based on *ab initio* lattice dynamics (LD).[5,6] However these methods do not include high-temperature effects and fail for dynamically stabilized structures with imaginary phonon frequencies. One solution to this is to extract vibrational frequencies from molecular dynamics (MD) data.[7–10]

For simple structures this is relatively straightforward, bcc titanium and zirconium being nice examples. In these materials the soft T_{1N} phonon eigenvector is well defined, and its frequency and width can be calculated from projection of the MD (or Monte Carlo) trajectories onto the relevant mode eigenvector, followed by Fourier transformation (FT).[8,11] For molecular systems IR spectra can be calculated from dipole contributions from each molecule.[12] Direct calculation of Raman spectra requires second-derivative (polarizability) calculations, which are still more challenging.

In lower-symmetry molecular systems there may be many modes which are formally Raman active, and the coupling between lattice and molecular modes is typically highly temperature dependent. Worst of all are the plastic crystal phases where the molecules can reorient in MD and the eigenvectors calculated from perturbation theory become totally irrelevant.

In this article we present a method for calculating Raman frequencies from MD and apply it to the particularly awkward and topical case of the high-frequency vibron modes in high-pressure hydrogen.

Although liquid and solid phases I and II of hydrogen have been well studied using MD,[13–15] much interest recently has focused on pressures around 200–300 GPa, where several phases are reported. Generally accepted are a low-temperature phase III (Ref. 16) and a high-temperature phase IV (Refs. 17 and 18). Theoretical predictions of many other phases[19,20] have been made, and Raman data suggest phase IV may itself have a subtle structural change at 270 GPa (Ref. 21).

At these pressures x-ray experiments are exceedingly difficult, while neutron diffraction is simply impossible; therefore most experimental data are extracted from Raman and infrared spectroscopy, alongside conductivity measurements. None of these techniques produces enough data to resolve crystal structures, so DFT studies have also been attempted.[18–20] Although these calculations typically ignore quantum effects on the protons, they still provide a useful indication of the likely structures. Very recent papers[22,23] applying path integral MD to high-pressure hydrogen show no qualitative behavioral change to the phase diagram: the main effect is that tunneling allows molecular rotations to occur at slightly lower temperatures than in classical MD, lowering the phase lines. Most importantly for the present work, the vibrational frequencies of the molecules are largely unchanged by the path integral dynamics.

Enthalpy is easily calculated in DFT, being the combination of total energy of binding of electrons to atoms plus the zero point energy (ZPE). Phase III should be the lowest enthalpy phase over a wide range of pressures. An *ab initio* structure search[20] unveiled a number of candidate phases with low total energy, and evaluation of normal modes, phonon frequencies, and ZPE gave rise to the prediction of a $C2/c$ symmetry phase. From the phonon calculation it is further possible to calculate polarization and polarizability, from which zero temperature Raman and infrared intensities may be deduced. These show reasonable agreement between $C2/c$ and experiment for phase III (Refs. 20 and 24–27).

Phase IV exists at higher temperatures and is therefore stabilized by entropy. High temperature calculations are more challenging for DFT. Using the quasiharmonic approach based on zero-temperature calculations, Pickard *et al.*[18] evaluated free energies at finite temperature to claim that phase IV should be a layered structure with alternating graphenelike hexagonal layers interspersed with ordered molecular layers and Pc symmetry. We refer to these layers as G type and Btype, respectively. The critical result here is that the two different layers give strongly Raman-active vibron modes at two very different frequencies. Experimental work also finds two vibron modes, lending support to a two-layer model.[17] The true nature of phase IV and its actual Raman spectrum can only reliably be calculated by including temperature effects, which is the subject of this paper.

II. THEORY

A. Finite temperature phonons

In order to study a particular vibrational mode in a crystal, we first define the calculation supercell and relax the structure at 0 K. The atoms are now located at positions given by $3N$ Cartesian coordinates X_j. We regard the supercell as a nonprimitive unit cell, in which case X_j are the basis positions.

We now do a lattice dynamics calculation at 0 K using either finite displacements[6] or DFPT.[5] This gives us a set of normal mode coordinates $\mathbf{e}_i$. With each of these normal modes we can calculate harmonic phonon frequency (ω_i), Raman activity, IR activity, and oscillator strength (R_i). i runs from 1 to $3N$, the number of normal modes. All of this is already standard in CASTEP.

From an MD simulation with T time steps we generate trajectories of the atoms, $x_i(t)$, at finite temperature. We can expand each Cartesian component of the trajectory in terms of the normal modes (ignoring translations):

$$x_j(t) = X_j + \sum_{i=4}^{3N} \alpha_i(t) e_{ij}.$$

So far all this is exact; we just made a linear transformation of the coordinate system. $\alpha_i(t)$ is fully determined by the MD. Similarly for velocities,

$$\dot{x}_j(t) = \sum_{i=4}^{3N} \dot{\alpha}_i(t) e_{ij}.$$

Now we assume that we are in the harmonic regime:

$$\alpha_i(t) = \mathrm{Re}[a_i \exp^{i(\omega_i t+\phi_i)}], \quad (1)$$

$$\dot{\alpha}_i(t) = \mathrm{Im}[a_i\omega_i \exp^{i(\omega_i t+\phi_i)}]. \quad (2)$$

This assumption means that a_i, ω_i, and ϕ_i are independent of time.

It is now straightforward to use the MD data to obtain ω_i by FT. The FT of α is problematic since at high temperature $\langle\alpha\rangle \neq 0$, but the same information is present in $\dot{\alpha}$ which is more convenient since $\langle\dot{\alpha}\rangle = 0$. In the harmonic limit, $FT[\dot{\alpha}_i(t)]$ is simply a δ function at $\omega = \omega_i$.

Note we have not used the frequencies from the lattice dynamics, we have reevaluated them from the MD. In the harmonic approximation, the same modes will be Raman/IR active in the MD as in the lattice dynamics. As usual, we can calculate the occupied phonon density of states from the velocity autocorrelation function:

$$FT\left[\sum_j \dot{x}_j(t)\dot{x}_j(0)\right] = FT\left[\sum_i\sum_j \dot{\alpha}_i(t)\dot{\alpha}_i(0)e_{ij}^2\right].$$

By analogy, the total Raman signal becomes

$$FT\left[\sum_{ij} R_i\dot{\alpha}_i(t)\dot{\alpha}_i(0)e_{ij}\right],$$

and we can obtain the mode frequency for each mode i from the peak in $FT\,[\dot{\alpha}_i(t)\dot{\alpha}_i(0)]$. In the harmonic limit, the Raman signal is simply the sum of the individual modes.

All of this has been applied in classical MD by numerous authors.[8] We now consider applying exactly the same process to an anharmonic MD. Modes with strong Raman/IR signals will still have strong Raman signals, since the polarizability ultimately depends on the motion of the atoms.

Care must be taken about the magnitude of the oscillations—in the harmonic case they will never equilibrate. To get close to equilibrium it seems sensible to set initial displacements and velocities based on temperature from the normal modes (with random phase ϕ):

$$\alpha_i(t=0) = \sqrt{kT/m\omega_i^2}\,\sin(\phi),$$

$$\dot{\alpha}_i(t=0) = \sqrt{kT/m}\,\cos(\phi).$$

This is done, e.g., in SCAILD,[12] but not automatically in CASTEP. This could be important in evaluating Raman intensities and line widths, since the anharmonic effects will depend on the phonon amplitude. However, for high-pressure studies the experimental Raman intensities depend strongly on the apparatus and are not used quantitatively.

B. Raman in dense H_2

In the particular case of hydrogen at high pressures and temperatures, the intense anharmonic proton motion makes the simple mode projection obsolete. Therefore we have to apply a more subtle method. For H_2 vibrons the Raman activity comes from the symmetric molecular stretch. It is therefore possible to associate α not with fixed normal modes but with the molecule stretches:

$$\alpha(t) = \sum_j \mathbf{r_j}(\mathbf{t}) - \mathbf{r_{j_m}}(\mathbf{t}),$$

where $\mathbf{r}_{j_m}(t)$ is the vector position of the molecular partner atom to j, at time t.

This requires us to identify molecules at each time step t (i.e., molecule labeling might change during the simulation, which could cause discontinuities in the velocity functions). Since the stretching modes can change at each time step, $\dot{\alpha}(t)$ is not the simple time derivative of $\alpha(t)$, but can be calculated by projecting the velocity $v_j(t)$ of each atom onto the stretching mode. Finally, the spectrum for the vibron modes is

$$FT\left\{\sum_j \mathbf{v_j}(\mathbf{t})\cdot[\mathbf{r_j}(\mathbf{t}) - \mathbf{r_{j_m}}(\mathbf{t})]\right\}.$$

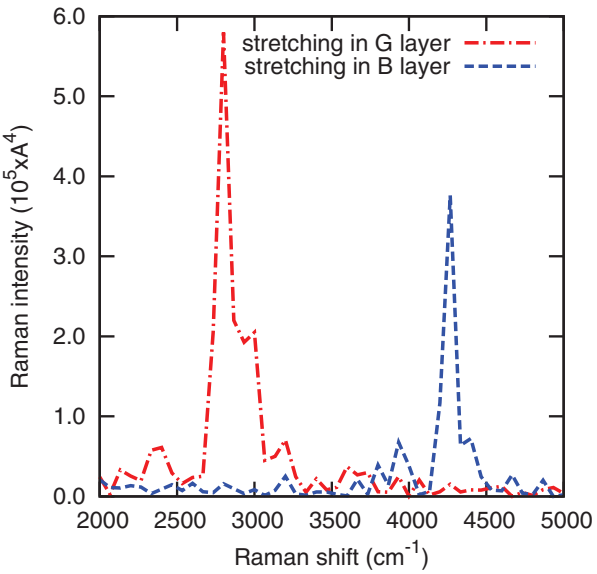

FIG. 1. (Color online) FT of projected velocity autocorrelation function of the two strongly Raman-active normal modes from MD calculations started in Pc at 250 GPa, 60 K. The lower-frequency mode (red, dash-dot) primarily involves in-phase molecular vibrations in the G layer, while the higher mode (blue, dashed) involves motion mainly in the B layer. They can be seen to be in excellent agreement with LD values 2845 cm^{-1} and 4187 cm^{-1}. The data in Fig. 6 were obtained from peaks in graphs such as this.

Here we investigate vibrons, but the method is completely general and could be extended to lower-frequency modes, provided that the Raman-active molecular pattern of oscillation can be identified.

Our DFPT-LD calculations show that all modes in the Pc phase are Raman active. To test our approach we first tried projecting the MD trajectories onto LD eigenmodes, multiplying by the calculated Raman intensity and taking the FT as shown in Fig. 1.

This method relies on the normal mode vectors being invariant over time and with temperature. At the lowest temperature, 60 K, the structure remains in the initial configuration (metastable Pc) and the projection method gives a Raman signal in precise agreement with LD, as it should for harmonic vibrations (reasons explained in Sec. II). At higher temperatures molecular rotations exchange atom positions, and this method fails completely. A bespoke choice of the projector e_{ij} is required.

Closer inspection shows that the strongly Raman-active vibron modes in all phases involve in-phase stretches of the H_2 molecules in the G and B layers. Therefore we make the ansatz that, independent of phase or molecular orientation, the Raman-active vibron modes will involve in-phase stretches. Extracting the Raman signal from the MD is now achieved by identifying molecular bond lengths at each step, which turns out to be always straightforward, and taking the time FT of the average projection of the velocity over the bond lengths. This method produces well-defined peaks not only in the low temperature harmonic regime but also in highly anharmonic high temperature phases with molecular rotations.

We note that this procedure requires well-defined molecules, but does not *require* us to identify layers. It is possible to identify the layers and to project only onto molecules in one layer. In that case each different layer produces a single peak: the higher-frequency peak being associated with the B layers. However, at pressure there is interference between layers,[28] so the sum of the layer projections does not give the correct Raman signal.

C. Finite size effects in dense H_2

Previous work, and our results shown later, indicate that solid H_2 is composed of well-defined layers, a secondary feature is the interaction between layers, and longer-ranged interactions, are weak. From the MD in phase IV, we note that adjacent B layers are not observed, while different possible translations of the G layers are observed, depending on starting configuration and kinetics: this implies little energy preference.

It is possible to understand the transition using a simple one-dimensional model Hamiltonian and to use this model to study when finite-size effects will make MD simulations unreliable. We assign differences in energies ($U_{GB} = U_G - U_B < 0$) and entropies ($S_{GB} = S_G - S_B < 0$) to each layer pair and a layer interaction term J_{ij}, where i and j label G or B layers:

$$H = \sum_{i=1}^{N} U_i - TS_i + J_{i,i+1}.$$

In this model the free energies of various N atom supercells with L layers are given in Table I.

Assuming $J_{GG} < J_{GB} < J_{BB}$ this model gives a phase diagram as shown in Fig. 2, including phases III and IV and a layered phase of free molecules (rather similar to

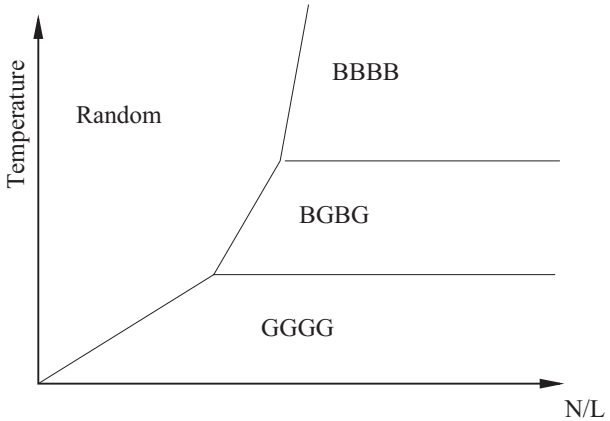

FIG. 2. Schematic drawing of the effect of finite size and temperature for the simple layer model. *Random* represents the dynamically changing finite-size phase observed in simulations with 24 or 48 molecules; $GGGG$, $BGBG$, and $BBBB$ are similar to phases III, IV, and I (or V), respectively. For small enough systems (low N/L) randomly oriented layers will always have the lowest free energy; however, in the thermodynamic limit $N/L \to \infty$ the phase sequence favors ordered layers. The actual values of T and N/L at the phase boundary depend on the parameters, which in turn depend on pressure.

TABLE I. Energies of four possible "phases" which could be realized in an $L = 4$ layer supercell, where in the general case L is the number of layers.

Phase	Stacking	Free energy
III	All G	$N(U_G - TS_G + J_{GG})$
I	All B	$N(U_B - TS_B + J_{BB})$
IV	$GBGB$	$N(U_B - TS_B + U_G - TS_G + J_{GB})/2$
	Random	$N(2U_B - 2TS_B + 2U_G - 2TS_G + J_{BB} + J_{GG} + 2J_{GB})/4 - LT\ln 2$

phase I and to theoretically predicted $Cmca$-4). Random stacking is favored by the $LT\ln 2$ term, which is significant only for small system sizes where the number of layers is comparable to the number of atoms. Consequently this is not a thermodynamically stable phase, but it may arise in finite-size simulations, e.g., in 48 molecule simulations each layer contains only 12 molecules. For a free energy difference of F_{BG} per atom between the distinct layer types, Boltzmann statistics shows that the probability of finding the B layer is $1/[1 + \exp(-12F_{BG}/KT)]$. Static calculation implies an $F_{BG} \approx 2$ meV/atom, so at 300 K the unfavored layer is present 28% of the time.

III. CALCULATIONS AND RESULTS

A. DFT and MD calculations

For our DFT calculations, we used the CASTEP code[29] with the Perdew-Burke-Ernzerhof (PBE) exchange-correlation functional which has become the standard for work in hydrogen.[30] Two different pseudopotentials were developed, an ultrasoft (300 eV cutoff) pseudopotential generated "on the fly"[31] for the molecular dynamics and a harder norm-conserving pseudopotential (1200 eV cutoff) for which Raman calculations are more easily carried out. The structural results obtained were similar for the two methods. For the DFPT lattice dynamics a single unit cell k-point set of $9 \times 5 \times 5$ was used, giving 69 independent k points. We find that MD simulations with 48 molecules are plagued with finite-size effects for reasons explained in the previous section, so all data here come from 288- or 768-atom calculations. The MD calculations are initiated in prerelaxed monoclinic supercells of either Pc ($\beta \approx 91°$) or $C2/c$ ($\beta \approx 144°$) structures. We started our simulations in NPT ensemble (constant number of particles [N], pressure [P] and temperature [T]), using a constant-stress Parrinello-Rahman barostat for equilibrating the structures in the right geometry. This initial stage was followed by longer NVE ensemble (constant number of particles [N], unit cell volume [V] and energy [E]), which were used for extracting the Raman spectra.

We have also extended Pickard's calculations of enthalpies across a wider range of pressures and with a variety of pseudopotentials (both ultrasoft and norm-conserving with various tunings) and exchange-correlation functionals. We find that these previous results are robust.[32] Despite qualitative agreement with experiment, the lower-frequency vibron in phase IV is observed at much higher frequency than calculated in DFPT (Sec. A of Ref. 33) and with very large width. We also used static calculations to investigate whether the G layers have atomic or molecular character. Mulliken bond analysis shows very clearly that the G-layer hexagonals are rings of three H_2 molecules: the Mulliken charge in the molecular bond is at least double that between molecules (1.5e vs 0.3e at 180 GPa, closing to 1.3e to 0.65e at 350 GPa). This result is consistent with snapshots from MD. The reduction in molecular fidelity with pressure is accompanied by a steady reduction in the band gap. The ZPE is large in these systems, and one might expect that the higher vibron frequency in the B layer would imply more ZPE associated with those atoms, perhaps driving segregation in HD mixtures. Surprisingly, when we projected the ZPE onto layers, we found that the difference was negligible and there was no driving force for segregation from ZPE.

B. Geometry at finite temperatures

We have conducted a range of MD simulations under various conditions of temperature and pressure, starting in either the Pc or the $C2/c$ phase as summarized in Sec. B of Ref. 33. Expect low kinetic barriers between the two phases, especially for small cells. We observe direct phase transformations between layered structures with and without the B-type free rotating molecular layer. The unit cells of Pc and $C2/c$ are sufficiently dissimilar that transformation between the two does not occur directly: a 53° change would be required to go from Pc to $C2/c$.

This makes it clear that MD alone cannot be relied upon to find the experimental structure: comparison with data is essential for validation.

1. MD starting with C2/c

In simulations starting with $C2/c$ we find a stable G-layered structure at temperatures up to around 300 K, above which a transformation occurs to a structure with alternating B and G layers similar to the monoclinic $C2$ (Ref. 20). The transformation mechanism is such that the $GBGB$-structure layers are formed almost orthogonal to the original, distorted $GGGG$ layers found in $C2/c$ (Fig. 3).

2. MD starting with Pc

For simulations starting with Pc, we find reversible transitions between two phases:

(i) a structure with $BGBG$ stacking similar to Pc, with threefold G-layer symmetry but with the B-layer molecules rotating about their centers,

(ii) a higher-temperature structure with hexagonal symmetry, stacked $BG'BG''$ with sixfold layer symmetry where the

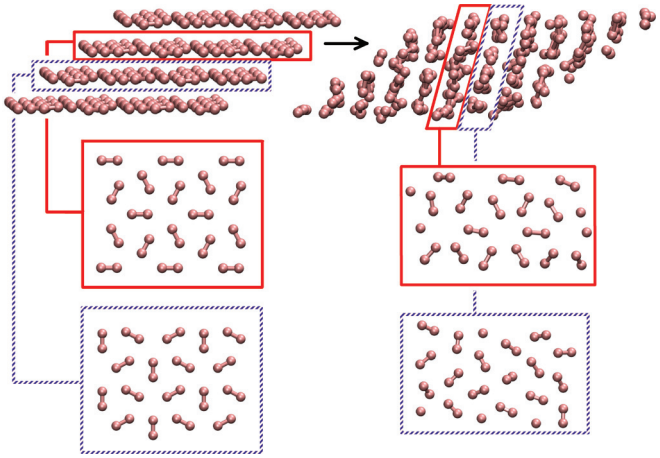

FIG. 3. (Color online) Transformation from phase III observed in NPT MD initialized in the $C2/c$ structure and heated at 250 GPa (in Ref. 33 check: simulation 28 from Sec. B and Fig. 7 from Sec. E). The upper image is the MD supercell, with layer projections shown below. Left: $C2/c$ (Ref. 19) relaxed at 250 GPa, with $GGGG$-layer stacking. Right: Snapshot from MD, after the transformation has occurred; alternating $BGBG$ stacking in the high-temperature phase (lower) is similar to that in phase IV, but there is some frustration, which leads to a lower Raman-mode frequency (around 2980 cm^{-1}).

G' layer has hexagonal symmetry and the G'' layer exhibits fast rebonding which enables rotation of the hexagonal motifs of three hydrogen molecules.

We refer to these phases as IVa and IVb (Fig. 4).

This transition is observed in two ways: a single MD run with a ramped temperature rise and subsequent decrease traverses a path IVa-IVb-IVa with little hysteresis [see simulation 27 from the Supplemental Material, Sec. B (Ref. 33)]; alternately, calculations at fixed T and P show the two phases [see the Supplemental Material, Sec. C (Ref. 33)]. We use these latter calculations as the basis for our Raman calculations to test the simulated structures against experiment.

For validating our results, we increased the cell size to 768 atoms, using 8 initial layers. We found the same transition from $BGBGBGBG$ to $BG'BG''BG'BG''$ at 270 GPa and room temperature,[33] supporting the hypothesis that the true nature of phase IV in hydrogen is best described by the $BG'BG''$ structure. It should be noted that our simulations have only four to eight layers, so more complex stackings may exist.

We extended the 768 atoms simulations to higher pressures to check the stability of phase IV. At 340 GPa, and room temperature, we found a phase transition to a structure that resembles the previously proposed $Cmca$-4 (Ref. 18) when average atomic positions are considered. We note the similarity between this $BBBB$ phase V and the molecular phases I and II; however, Mulliken bond-charge analysis suggests that the bond charge in phase V is much reduced. The transition to $Cmca$-4 was also found by Liu *et al.*[34] in their metadynamics calculations. We conclude that the already-achievable experimental conditions[17] are already close to the transition to a $Cmca$-4 "phase V" state.

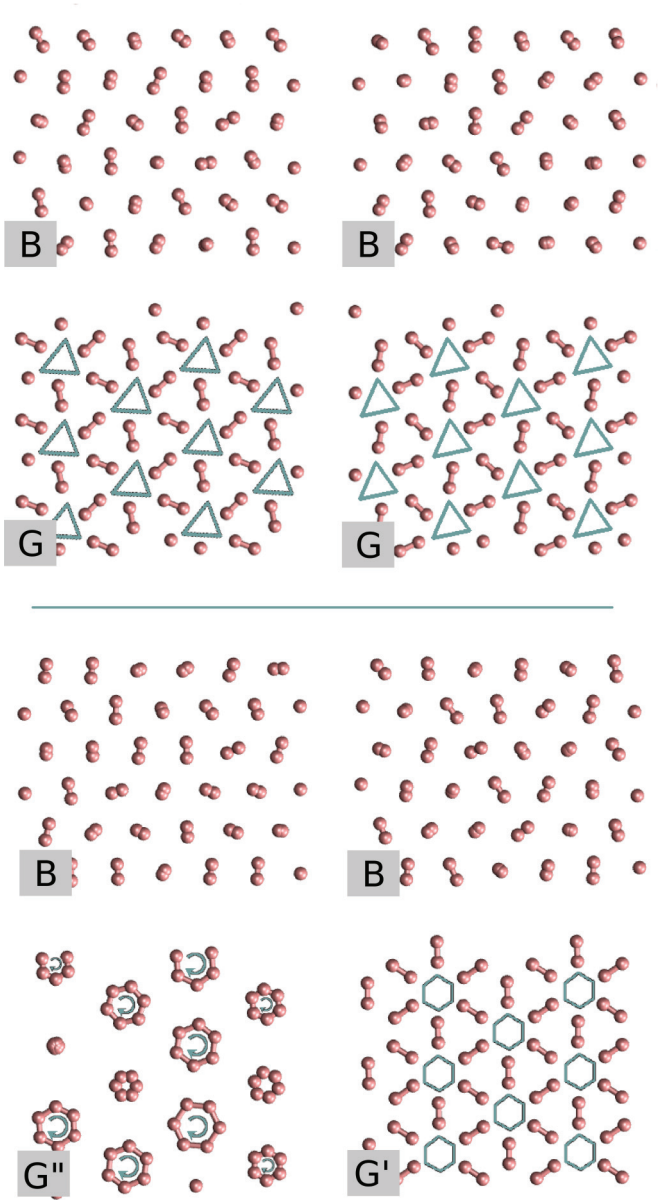

FIG. 4. (Color online) Time-averaged (1 ps) atomic positions from simulations at pressure/temperature/initial configuration. Top: Phase IVa, stacked $BGBG$, 220 GPa, 220 K, Pc. Bottom: Phase IVb, stacked $BG''BG'$, 270 GPa, 220 K, Pc. The time averaging was chosen large enough to capture the rotation of the B-layer molecules and G''-layer motifs, making these smaller. After a long time and complete rotations, the motifs become degenerate points at the central position representing 2 B and 6 G'' atoms, respectively. The G'-layer image is time invariant. Note the strong sixfold symmetry in phase IVb compared with IVa. Many other similar figures are given in the Supplemental Material,[33] showing the trends with P and T.

C. Raman spectra at finite temperature

1. Phase III

In Fig. 5 we show the pressure dependence of the MD-calculated Raman vibron frequencies compared with the experimental data. In phase III the MD has spectacular agreement with experiment.[24] DFPT gives similar frequencies, but with a much lower slope (Fig. 6). The simulation which

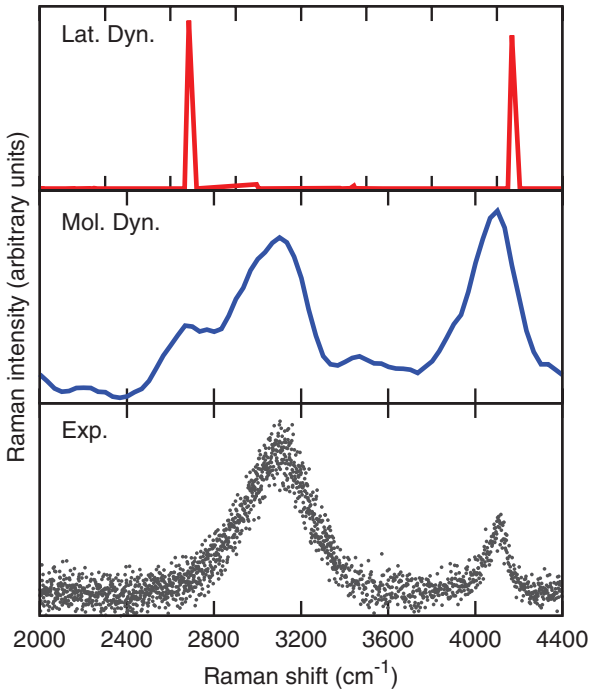

FIG. 5. (Color online) Comparison of Raman spectra at 270 GPa. Top: DFPT lattice dynamics calculation (0 K). Middle: MD calculation of Raman signal from phase IVb at 220 K. Bottom: Experimental data at room temperature from Howie *et al.*[17] Note that the sensitivity of the detector is reduced at high frequency,[21] so the peak amplitudes are not directly comparable.

started with $C2/c$ at 250 GPa/220 K was also driven through the phase transition by heating to 300 K, where it formed a frustrated monoclinic, $GBGB$-stacked structure as shown in Fig. 3, leading to the appearance of a second vibron. This structure has a mixed-layer character like phase IV, but the frequencies are clearly not in agreement with experiment (inset to Fig. 6). This metastable structure illustrates the hierarchical natural of the bonding: a primary tendency to form layers, a secondary tendency to order as B or G within the layers, and a tertiary effect of interlayer interactions. A similar transformation, between *po-hcp* (phase IV) and $C2/c$ (phase III) has been seen by Liu *et al.*[34]

2. *Phase IV*

Figure 6 also shows that the phase IVb structure with rotating trimers does correctly reproduce the experimental frequencies, while the phase IVa and Pc structures do not. The phase IVb lower peak seems to be composed of two overlapping peaks as shown in Fig. 5. We can project the symmetric stretch modes layer by layer: for phase IVb with three different layers, this gives three different frequencies. The near-equal strength of the G' and G'' peaks is due to having one layer of each in the simulation. This ratio is determined by finite size, and in reality the lower peak may be weaker. In fact, this feature is also probably present as a shoulder in the experiment, although its effect could be interpreted as an extended peak width (see Fig. 6). We therefore identify our phase IVb with the experimentally observed phase. Phase IVa

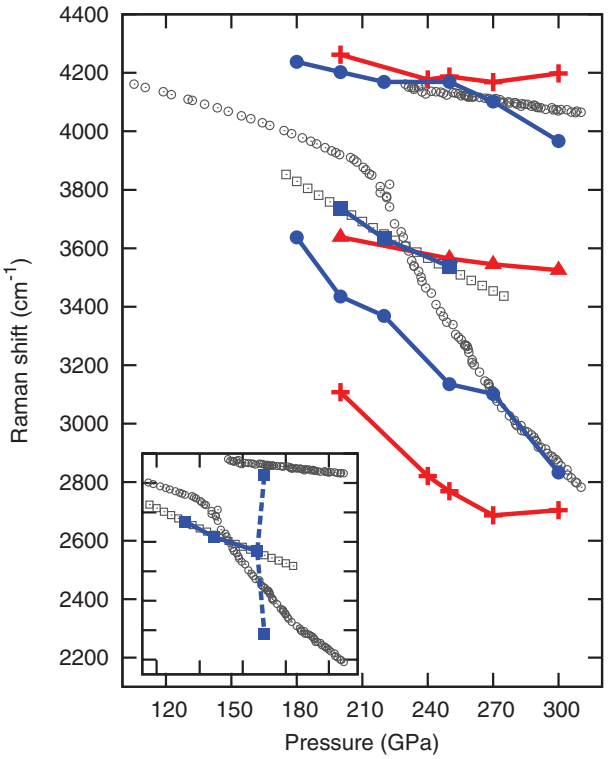

FIG. 6. (Color online) Pressure dependence of the vibron peaks. Experiment (gray): Phase IV, 300 K (Ref. 17) (open circles); phase III, 90 K (Ref. 24) (open squares). MD (blue): Initialized in Pc at 220 K (solid circles) and in $C2/c$ at 220 K (solid squares). Lattice dynamics (red): In Pc (crosses) and in $C2/c$ (solid triangles). The discontinuity in the Pc MD at 270 GPa corresponds to the IVa-IVb transition. Inset: The bifurcation of the $C2/c$ MD at 250 GPa corresponds to the two Raman peaks which were calculated after $C2/c$ was heated to 300 K and it transformed to a distorted structure similar to Pickard's $C2$.

occurs in a region of pressure-temperature (PT) space occupied by phase III, so we regard it as metastable.

IV. DISCUSSION

The analysis of the detailed molecular motions gives a clear, intuitive picture of the high pressure phase behavior of hydrogen. At low temperature we observe a series of G layers to be the stable structure. Above 60 K, molecular motion means that this layer has threefold symmetry, but at 0 K we found symmetry-breaking distortion gives the $C2/c$ structure. Whether this distortion could be observed once quantum effects of the protons are considered is unclear—path integral MD calculations would be needed to determine this.

At a temperature of around 250–300 K (depending on pressure[33]) the transformation to phase IV (our phase IVb) occurs. The explanation for the entropy-driven transition is evident in the rapid rotational movement of the B-layer atoms. These rotations provide the primary entropy difference between phases III and IV. The same rotation in the B layer means that time-averaged molecular positions have hexagonal symmetry. At higher temperature, the fast rebonding in the G'' layers enables the rotation of the hexagonal motifs and adds to the entropy, stabilizing the phase IVb structure. Once this rotation begins, the nonrotating G layer adopts hexagonal

symmetry. At lower temperatures, no rotation occurs and the hexagonal motifs are distorted and symmetry broken, as in Pc.

Our 768- and 288-atom results are significantly different from previous MD work[34,35] which has been equivocal about the structure of phase IV, due to reorienting of molecules, transformation of one layer type to another, and "mixed" phases of apparently random B- and G-layer stacking. In our own calculations with 24 or 48 molecules per unit cell we find the same behavior as described in previous work.[34,35] These finite-size effects could be responsible for apparent rapid diffusion.

Proton "diffusion" in G layers requires two distinct steps:

(i) Rebonding within the rotating hexagonal motifs, a local process which can contribute to Raman broadening and may be enhanced by tunneling. This is seen in all simulation sizes at sufficient temperature and is illustrated in Fig. 4 (see also Sec. D of Ref. 33).

(ii) Rearrangement of the motifs themselves, a process which must occur systemwide. We see this frequently in 48 atom simulations, very rarely with 288 and never with 768. The rebonding and rotation of these motifs was also described by Liu *et al.*,[34] who also showed large diffusion of hydrogen which implies the rearrangement step. The rotation appears in all our simulations and is responsible for the broadening of the lower vibron; however, we regard the reconstruction as a finite-size effect and do not expect rapid diffusion in phase IV (Refs. 36 and 37).

Our phase IVb structure is the best model for the observed phase IV. Our molecule-based technique shows that vibron peaks appear which can be associated with each layer type. The B layer gives the highest frequency, and analysis of the Raman-active modes associated with G' and G'' layers in phase IVb gives two distinct vibrons of slightly different frequency. In the overall pattern, these peaks overlap to give a single peak with a shoulder. The wide variety of environments in which the G-layer molecules find themselves leads to a very broad Raman width. The B-layer molecules are well defined and the Raman peak associated with them is sharper. In Pc, phases IVa and IVb, the B-layer vibron has similar frequency; however the G layer modes are quite different.

V. CONCLUSIONS

In sum, we have shown how Raman frequencies can be extracted from MD data allowing direct comparison to experiment. We have applied the method to hydrogen at high pressure, showing that the anharmonicity is so extreme as to invalidate the use of DFPT normal modes, but that in-phase vibrons are the appropriate coordinates for projection.

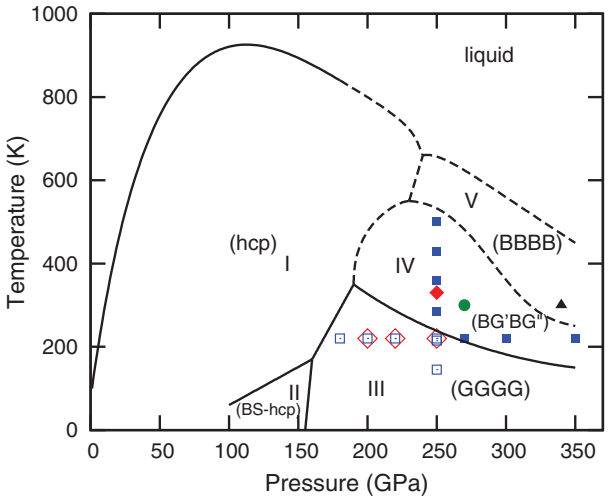

FIG. 7. (Color online) Proposed phase diagram of dense hydrogen. Solid lines are from previous work.[21,34] The dashed lines are inferred from our simulations, which are labeled as follows (started in, transformed to): $C2/c$ that remained a $GGGG$ stacking (open red diamonds); $C2/c$ that transformed to a $BGBG$ stacking similar to Pickard's $C2$ (solid red diamond); Pc that stayed in a $BGBG$ stacking and we regard as metastable (open blue squares); Pc that transformed to a $BG'BG''$ stacking (solid blue squares and solid green circle); and finally the simulation that started with Pc and changed to a $BBBB$ stacking similar to Pickard's $Cmca$-4 (solid black triangle). Note that both the solid green circle and the solid black triangle labels depict simulations containing 768 atoms, while all the other labels illustrate simulations with 288 atoms.

Our simulations show several different structures, some of which are doubtless metastable, but by comparison to experiment we identify phase III with a structure similar to $C2/c$, phase IV with a high-entropy hexagonal structure of rotating molecules and trimer motifs. We also predict the transition to a state similar to the proposed $Cmca$-4 close to the experimental conditions that have already been achieved [320 GPa (Ref. 17)]. In Fig.7, we propose an approximate phase diagram, which best concludes our observations. Our simulations give no support to the notion that phase IV exhibits long-range proton transfer, proton tunneling, or mixed molecular and atomic character.

ACKNOWLEDGMENT

We acknowledge E. Gregoryanz for numerous useful discussions and EPSRC for a studentship (I.B.M.).

[1]F. A. Gorelli, S. F. Elatresh, C. L. Guillaume, M. Marqués, G. J. Ackland, M. Santoro, S. A. Bonev, and E. Gregoryanz, Phys. Rev. Lett. **108**, 055501 (2012).

[2]M. Marqués, M. I. McMahon, E. Gregoryanz, M. Hanfland, C. L. Guillaume, C. J. Pickard, G. J. Ackland, and R. J. Nelmes, Phys. Rev. Lett. **106**, 095502 (2011).

[3]D. Porezag and M. R. Pederson, Phys. Rev. B **54**, 7830 (1996).

[4]S. Baroni, S. de Gironcoli, A. Dal Corso, and P. Giannozzi, Rev. Mod. Phys. **73**, 515 (2001).

[5]K. Refson, P. R. Tulip, and S. J. Clark, Phys. Rev. B **73**, 155114 (2006).

[6]G. J. Ackland, M. C. Warren, and S. J. Clark, J. Phys.: Condens. Matter **9**, 7861 (1997).

[7]C. Z. Wang, C. T. Chan, and K. M. Ho, Phys. Rev. B **42**, 11276 (1990).

[8]U. Pinsook and G. J. Ackland, Phys. Rev. B **59**, 13642 (1999).
[9]J. A. Thomas, J. E. Turney, R. M. Iutzi, C. H. Amon, and A. J. H. McGaughey, Phys. Rev. B **81**, 081411(R) (2010).
[10]T. Lan, X. Tang, and B. Fultz, Phys. Rev. B **85**, 094305 (2012).
[11]P. Souvatzis, S. Arapan, O. Eriksson, and M. Katsnelson, Europhys. Lett. **96**, 66006 (2011).
[12]J. Sun, D. Bousquet, H. Forbert, and D. Marx, J. Chem. Phys. **133**, 114508 (2010).
[13]S. Scandalo, Proc. Natl. Acad. Sci. USA **100**, 3051 (2003).
[14]J. Kohanoff, S. Scandolo, G. L. Chiarotti, and E. Tosatti, Phys. Rev. Lett. **78**, 2783 (1997).
[15]J. Kohanoff, S. Scandolo, S. de Gironcoli, and E. Tosatti, Phys. Rev. Lett. **83**, 4097 (1999).
[16]H. K. Mao and R. J. Hemley, Rev. Mod. Phys. **66**, 671 (1994).
[17]R. T. Howie, C. L. Guillaume, T. Scheler, A. F. Goncharov, and E. Gregoryanz, Phys. Rev. Lett. **108**, 125501 (2012).
[18]C. J. Pickard, M. Martinez-Canales, and R. J. Needs, Phys. Rev. B **85**, 214114 (2012); **86**, 059902(E) (2012).
[19]C. J. Pickard and R. J. Needs, Phys. Rev. Lett. **97**, 045504 (2006).
[20]C. J. Pickard and R. J. Needs, Nat. Phys. **3**, 473 (2007).
[21]R. T. Howie, T. Scheler, C. L. Guillaume, and E. Gregoryanz, Phys. Rev. B **86**, 214104 (2012).
[22]X.-Z. Li, B. Walker, M. I. J. Probert, C. J. Pickard, R. J. Needs, and A. Michaelides, J. Phys.: Condens. Matter **25**, 085402 (2013).
[23]M. A. Morales, J. M. McMahon, C. Pierleoni, and D. M. Ceperley, Phys. Rev. Lett. **110**, 065702 (2013).
[24]Y. Akahama, H. Kawamura, N. Hirao, Y. Ohishi, and K. Takemura, J. Phys.: Conf. Ser. **215**, 012056 (2010).
[25]Y. Akahama, M. Nishimura, H. Kawamura, N. Hirao, Y. Ohishi, and K. Takemura, Phys. Rev. B **82**, 060101(R) (2010).
[26]P. Loubeyre, F. Occelli, and R. Letoullec, Nature (London) **416**, 613 (2002).
[27]C.-S. Zha, Z. Liu, and R. J. Hemley, Phys. Rev. Lett. **108**, 146402 (2012).
[28]H. C. Hsueh, M. C. Warren, H. Vass, G. J. Ackland, S. J. Clark, and J. Crain, Phys. Rev. B **53**, 14806 (1996).
[29]M. D. Segall, P. J. D. Lindan, M. J. Probert, C. J. Pickard, P. J. Hasnip, S. J. Clark, and M. C. Payne, J. Phys.: Condens. Matter **14**, 2717 (2002).
[30]J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).
[31]Using the input string 'H 0|0.7|2|6|8|10L(qc = 10)' in the CASTEP parameter file.
[32]We note that convergence of the Raman signal requires a much higher k-point sampling than either structure or phonon calculations.
[33]See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevB.87.174110 for additional illustrations and technical details.
[34]H. Liu and Y. Ma, Phys. Rev. Lett. **110**, 025903 (2013); H. Liu, L. Zhu, W. Cui, and Y. Ma, J. Chem. Phys. **137**, 074501 (2012).
[35]A. F. Goncharov, J. S. Tse, H. Wang, J. Yang, V. V. Struzhkin, R. T. Howie, and E. Gregoryanz, Phys. Rev. B **87**, 024101 (2013).
[36]M. A. Morales, C. Pierleoni, E. Schwegler, and D. M. Ceperley, Proc. Natl. Acad. Sci. USA **107**, 12 (2010).
[37]W. Lorenzen, B. Holst, and R. Redmer, Phys. Rev. B **82**, 195107 (2010).